جامعة 8 ماي 1945

كلية العلوم الإنسانية و الاجتماعية.

قسم علم النفس.

ماستر 1 علم النفس العيادي.

المقياس : الإحصاء و تحليل المعطيات. الأستاذ: د.مكناسي محمد

الفصل الثاني:أساسيات الإحصاء الوصفي

- تنظيم و عرض البيانات الإحصائية.
- مقاييس النزعة المركزية و التشتت.
- التوزيع الاعتدالي.

1-تنظيم و عرض البيانات الإحصائية

مهما يكن هدف الدراسة و أداة جمع البيانات ، فان البيانات في الدراسات السيكولوجية تكون عادة كثيرة ، لهذا فالباحث لا يمكنه الاحتفاظ بها في حالتها الخام التي تتطلب صفحات عديدة و تصعب عمليات التحليل اللاحقة لدى على الباحث تنظيمها بإحدى الطرق الآتية: الجداول الإحصائية و الرسوم البيانية أو الطريقتين معا. إن الغرض الأساسي من عملية تنظيم البيانات هو محاولة الاستفادة و الخروج بملامح عامة من هذه البيانات ، لان البيانات الخام لا نستطيع الاستفادة منها بشيء إلا عندما تنظم ، حيث يسعى الباحث في هذه الخطوة إلى تقديمها بشكل مبسط لتسهل عملية معالجتها وتفسيرها، و البيانات أو الدرجات الخام هي الدرجات التي نحصل عليها مباشرة من تطبيق الاختبارات أو أدوات القياس و لم تجر عليها أية عمليات إحصائية.

تنظيم البيانات ← في صورة جداول (الجدول التكراري البسيط - ج .ت. للفئات- ج.ت.المزدوج)

تنظيم البيانات ← في صورة رسوم بيانية(المنحنى التكراري-المضلع التكراري-المدرج التكراري).

- **العرض الجدولي للبيانات**: يسهل الجدول تمثيل كميات كبيرة من البيانات بطريقة جذابة وسهلة القراءة ومنظمة ، ويتم تنظيم البيانات في صفوف وأعمدة ، هذا هو أحد أكثر أشكال عرض البيانات استخدامًا لأن جداول البيانات سهلة الإنشاء والقراءة.

1- جدول تفريغ البيانات

- الباحث بعد جمعه للبيانات يقوم بتفريغها في مصفوفة تسمى مصفوفة البيانات ، و هي عبارة عن جدول به عمودا يخصص لأرقام الاستمارات (أفراد عينة الدراسة). و صفا به أسماء المتغيرات (بيانات شخصية+ متغير الدراسة1 +متغير الدراسة2+.....ثم صفا أخر يتضمن تفاصيل كل متغير). كما هو مبين في الجدول أسفله: (و هو مشابه لصفحة التفريغ في برنامج **spss.**).

المتغيرات	بيانات شخصية			متغير الدراسة1					متغير الدراسة2						
الاستمارات (الافراد)	جنس	سن	ح.ع	ع	ع2	ع3	ع4	ع5	ع1	ع2	ع3	ع4	ع5	.	.
1	ذ	20	م	4	1	2	5	3	2	5	4	1	3		
2	أ	22	ع	1	1	2	3	1	5	2	2	3	4		
3															
4															
5															
6															
7															
.															
.															
.															

2- العرض الجدولي للبيانات النوعية

- **طريقة الجداول التكرارية:** غالبا ما يتكون الجدول في هذه الحالة من عمودين، حيث يمثل العمود الأول الصفات المختلفة للمتغير النوعي ونرمز لها بـ **(x_i)**، والعمود الثاني يمثل تكرار كيفية الصفة والذي نرمز له بـ **(n_i)**، ويرسم الجدول بصفة عامة على النحو التالي:

المتغير (x_i)	التكرار المطلق (n_i)
X_1	n_1
X_2	n_2
X_3	n_3
.	.
.	.
X_k	n_k

ويمكننا أن نجد معلومات إضافية في الجدول، أي تكرارات أخرى مثل التكرار النسبي، حيث يتم حسابه بقسمة تكرار كل صفة على مجموع التكرارات، ويرمز له بـ **f_i** ، ويحسب بالطريقة التالية:

$$\text{التكرار النسبي} = \frac{\text{تكرار الصفة}}{\text{مجموع التكرارات}} \qquad f_i = n_i / n_k$$

مثال: البيانات التالية تمثل تقدير عشرة طلبة في مادة الإحصاء الوصفي

جيد، ممتاز، متوسط، جيد، متوسط، حسن، ممتاز، متوسط، جيد، حسن،

المطلوب: قم بإعداد جدول توزيع تكراري:

الحل:

التقديرات (x_i)	عدد الطلبة (n_i)
ممتاز	02
جيد	03
حسن	02
متوسط	03
المجموع	10

ومن خلال هذا العرض والتنظيم البسيط في الجدول يتسنى لنا من النظرة الأولى معرفة توزيع الصفة بوضوح، أفضل مما كانت عليه في البداية، وهنا تظهر أهمية العرض الجدولي.

3-العرض الجدولي للبيانات الكمية

أ- في حالة متغير كمي منفصل:

المتغيرات الكمية المنفصلة هي التي تأخذ بياناتها أرقاما عددية صحيحة مثل عدد الأولاد أو عدد المعلمين وغيرها، ويتم تنظيم هذا النوع من المتغيرات في جدول به عمودين، حيث نصنف في العمود الأول قيم الظاهرة (x_i) وفي العمود الثاني تكرارات هذه القيم.

مثال[1]: البيانات التالية تمثل عدد الأفراد في عينة مكونة من 30 أسرة:

5	4	3	4	2	2	5	4	4	2	5	4	2	3	2
3	5	4	3	5	4	3	5	4	5	4	3	4	5	3

المطلوب: عرض البيانات في جدول توزيع تكراري.

الحل:

حجم الأسرة (x_i)	عدد الأسر (التكرارات n_i)
2	5
3	7
4	10
5	8
المجموع	30

ب- في حالة متغير كمي متصل

في حالة المتغير الكمي المتصل تنظم البيانات وفق نوعين من الجداول، وهذا حسب عدد أفراد العينة، فإذا كان أفراد العينة مساو لـ (30) أو أقل فإن الجدول ينظم مثلما هو الحال في المتغير الكمي المنفصل، اما إذا كان عدد أفراد العينة أكبر من (30)، فإن الجدول ينظم عن طريق فئات، ونتعرف فيما يلي على كيفية تنظيم مثل هكذا جداول بالتطبيق مباشرة من خلال المثال التالي.

مثال: فيما يلي درجات 39 طالبا في اختبار الإحصاء.

12	13	7	6	8	5	4	7	10	7
9	13	12	15	9	11	12	11	9	2
17	8	13	3	14	9	3	19	14	5
	10	15	17	12	13	14	9	11	7

المطلوب:1/ حدد المتغير المدروس و نوعه ؟2/ كون جدول تكراري للبيانات السابقة ؟

الحل:

✓ **تحديد المدى المطلق = أكبر قيمة – أصغر قيمة**

أكبر قيمة: 19 أصغر قيمة: 2 المدى = 19 – 2 = **17**

✓ نحدد عدد الفئات: ويمكن تحديد عدد الفئات وفق القانون التالي:

عدد الفئات= 1 + (3.3 X اللوغاريتم العشري لحجم العينة) = 1 + (3.3 X 1.59) ≈ **6**

✓ نحدد **طول الفئة = المدى / عدد الفئات** = 17 / 6 ≈ **3**

الآن توفرت لدينا كل المعلومات لتنظيم القيم في شكل فئات، مع الأخذ بعين الاعتبار أن أول فئة يمكن أن تبدأ بأصغر قيمة في التوزيع.

وتكتب الفئة على النحو التالي: [2-5[حيث (2) يسمى الحد الأدنى للفئة، أما (5) فهو الحد الأعلى، وتقرأ هذه الفئة كالتالي: من 2 إلى أقل من خمسة، أي عندما نريد حساب التكرارات لا نقوم عد (5) في هذه الفئة وإنما في الفئة الثانية. وفي مثل هذا النوع من الجداول يهمنا كثيرا أن نعرف مراكز الفئات، حيث تحسب هذه الأخيرة بالقانون التالي: **مركز الفئة = الحد الأدنى + الحد الأعلى / 2**

ومعرفة مركز الفئة يساعدنا كثيرا في التمثيل البياني للمعطيات كما سنوضح.

الفئات	التكرارات	مركز الفئة	التكرار النسبي	التكرار المئوي
[2-5[	4	3,5	0.10	10
[5-8[	7	6,5	0.18	18
[8-11[	8	9,5	0.20	20
[11-14[	12	12,5	0.31	31
[14-17[	5	15,5	0.13	13
[17-20[	3	18,5	0.08	8
المجموع	39		01	100

4- الجدول المزدوج : وهو الجدول الذي يربط بين متغيرين في نفس الوقت وكل متغير منهم له فئاته فيتم بناؤه بإتباع عدة خطوات هي :

1- تحديد المتغيرين. 2- تحديد المتغير المستقل والمتغير التابع . 3- تحديد فئات كل من المتغيرين.

4- تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أي يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل أي يكون عمودياً.

5- وضع العلامات التي تمثل التكرار. 6- إعادة كتابة الجدول بالأرقام .

مـثـال : الجدول التالي يوضح البيانات التي حصل باحث في دراسة بين النوع و مشاهدة البرامج التعليمية لمجموعة من طلاب الصف الثالث الثانوي على النحو التالي :

النوع	مشاهدة البرامج	النوع	مشاهدة البرامج
ذكر	يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	لا يشاهد
أنثى	يشاهد	أنثى	لا يشاهد
ذكر	لا يشاهد	أنثى	يشاهد
أنثى	يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	لا يشاهد
ذكر	لا يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	يشاهد
أنثى	لا يشاهد	أنثى	لا يشاهد

والمطلوب تكوين الجدول المزدوج للعلاقة بين المتغيرين (النوع ومشاهدة البرامج التعليمية) ؟

الـحل :

1- المتغيرين (النوع – مشاهدة البرامج التعليمية)

2- المتغير المستقل هو النوع والمتغير التابع هو مشاهدة البرامج التعليمية .

3- فئات المتغير النوع هي (ذكور – إناث)

فئات المتغير مشاهدة البرامج التعليمية (يشاهد – لا يشاهد)

4- تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أى يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل أي يكون عمودياً .

كالتالى :

النوع / المشاهدة	ذكور	إناث
يشاهد		
لا يشاهد		

5- وضع العلامات .

النوع / المشاهدة	ذكور	إناث
يشاهد	/////	////
لا يشاهد	/////	//////

6- إعادة كتابة الجدول بالأرقام .

النوع / المشاهدة	ذكور	إناث	المجموع
يشاهد	5	4	9
لا يشاهد	5	6	11
المجموع	10	10	20

- **العرض البياني للبيانات:**

1-العرض البياني للبيانات النوعية:

يعتبر العرض البياني كذلك طريقة لعرض البيانات الوصفية ونجد من أهمها:

- ✓ **الدائرة النسبية**: وهي عبارة عن دائرة مقسمة إلى أجزاء أو قطاعات، بحيث كل جزء يمثل خاصية من خصائص المتغير المدروس، بحيث يتم حساب زاوية كل خاصية بالقانون التالي:

زاوية القطاع = تكرار الخاصية / مجموع التكرارات × 360° ويجب أن يكون مجموع زوايا القطاع 360°

مـثـال : الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية العلوم الاجتماعية و الإنسانية جامعة قالمة والمطلوب عرض هذه البيانات باستخدام طريقة الدائرة البيانية ؟

القسم	التاريخ	الاجتماع	الإعلام	علم النفس	الفلسفة
عدد الطلاب	650	500	400	350	550

الـحل :

نحسب مجموع التكرارات = 650+500+400+350+550 مجموع التكرارات = 2450

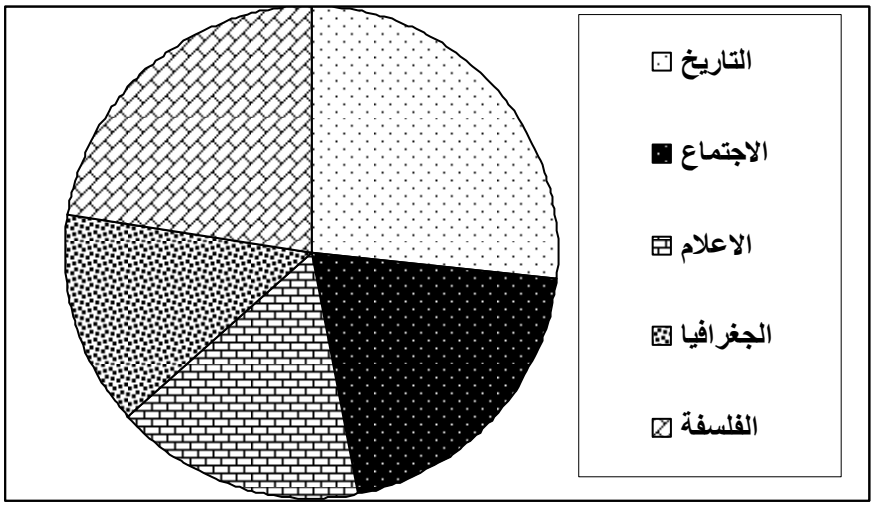
زاوية قطاع التاريخ = $\frac{650}{2450} \times 360 = 95.5^0$

زاوية قطاع الاجتماع = $\frac{500}{2450} \times 360 = 73.5^0$

زاوية قطاع الإعلام = $\frac{400}{2450} \times 360 = 58.7^0$

زاوية قطاع الجغرافيا = $\frac{350}{2450} \times 360 = 51.4^0$

زاوية قطاع الفلسفة = $\frac{550}{2450} \times 360 = 80.8^0$

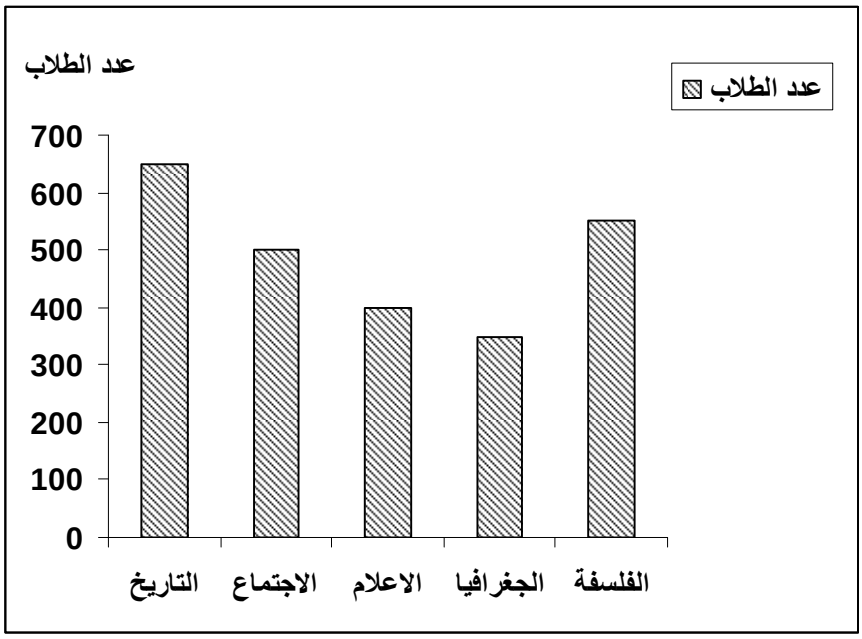


- ✓ **الأعمدة البيانية** يعتمد هذا التمثيل البياني أساسا على رسم معلم، نضع فيه على المحور الأفقي كيفيات الصفة وعلى المحور العمودي التكرارات المطلقة أو النسبية، ونرسم مستطيلا قاعدته على المحور الأفقي ممثلة بواسطة الكيفية ويتناسب طردا مع التكرارات، وتكون قاعدة المستطيلات ثابتة.

مـثـال : الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية العلوم الاجتماعية و الإنسانية جامعة قالمة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية البسيطة ؟

القسم	التاريخ	علم الاجتماع	الإعلام	علم النفس	الفلسفة
عدد الطلاب	650	500	400	350	550

الحل: توزيع الطلبة حسب التخصص.



1-العرض البياني للبيانات الكمية المبوبة:

يعتبر العرض البياني شكلا من أشكال تلخيص البيانات وفي الحقيقة يعطينا معلومات عن البيانات أسرع من الجداول .والمقصود بالبيانات المبوبة تلك البيانات المقسمة إلى فئات وهناك عدة طرق لعرض البيانات المبوبة. ، أما عن الرسومات البيانية الخاصة بالمتغيرات الكمية المبوبة فهي: المدرج التكراري، المنحنى التكراري والمضلع التكراري.

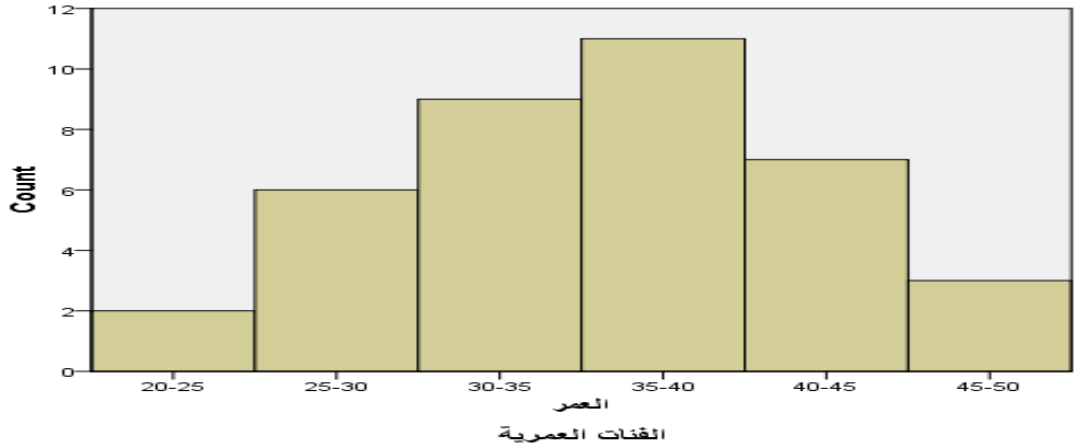
- **المدرج التكراري :** أحد طرق عرض البيانات المبوبة حيث يتم تخصيص عمود لكل فئة وتكرارها ، بحيث يكون طول الفئة هي قاعدة العمود والتكرار هو ارتفاع العمود ، ، أما بالنسبة لمنتصف العمود فيكون هو مركز الفئة .

مـثـال : اعرض هذا الجدول بيانياً باستخدام المدرج التكراري ؟

فئات العمر	20-	25-	30-	35-	40-	45-
عدد العمال	2	6	9	11	7	3

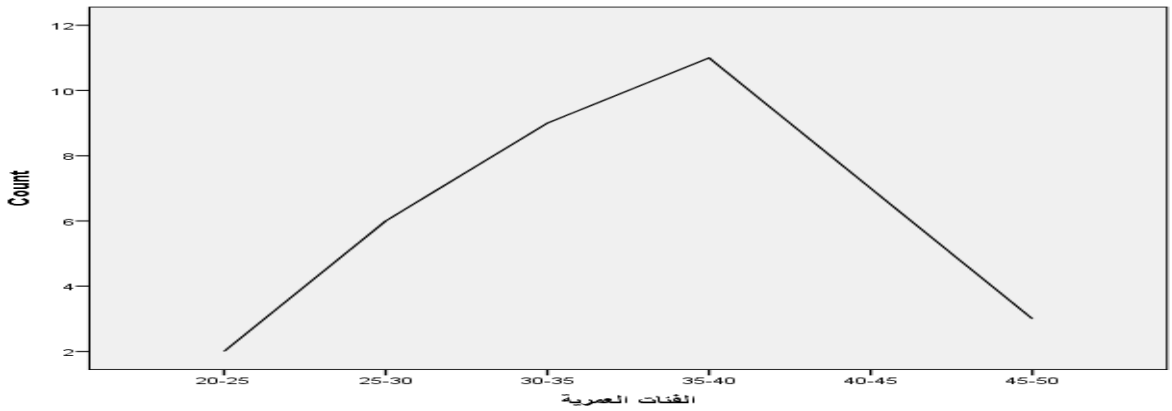
الـحل:

الفئات	ك	مركز الفئة
20-	2	22.5
25-	6	27.5
30-	9	32.5
35-	11	37.5
40-	7	42.5
45-	3	47.5



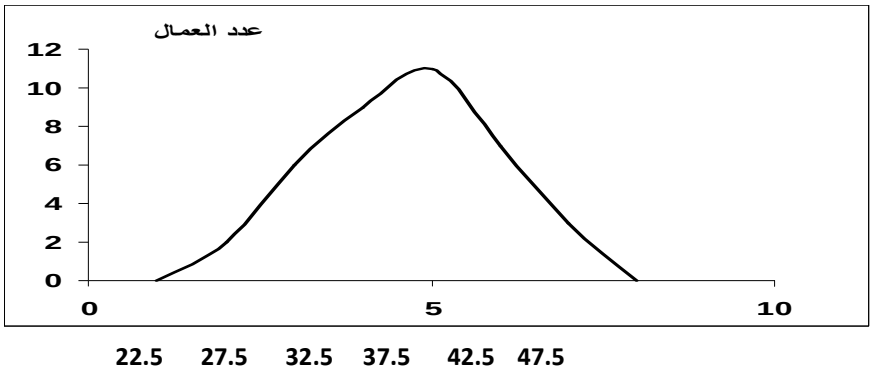
❖ **المضلع التكراري :** تخصص لكل فئة وتكرارها نقطة ، بحيث يكون الاحداثى السيني لها هو مركز الفئة بينما الاحداثى الصادي لها هو التكرار ، ثم نوصل كل نقطتين متتاليتين بخط مستقيم بالمسطرة .

نفس المثال السابق.:



❖ **المنحنى التكراري :** بعد رصد النقاط كما فى الطريقة السابقة نوصل كل نقطتين متتاليتين بمنحنى باليد .

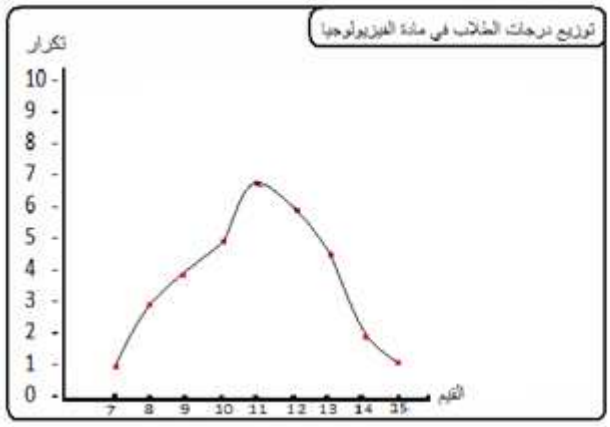
نفس المثال السابق.:



مثال: إليك درجات 33 طالب في مقياس الإحصاء.

القيم	7	8	9	10	11	12	13	14	15
التكرار	1	3	4	5	7	6	4	2	1

في البداية نرسم تقاطع محور أفقي يمثل القيم مع محور عمودي يمثل تكرارات تلك القيم، ثم نحدد نقاط تقاطع كل قيمة مع تكرارها ثم نربط تلك النقاط التي تم تحديدها لنتحصل على الشكل التالي:



2-مقاييس النزعة المركزية و مقاييس التشتت.

ان الهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.ومن أهم مقاييس النزعة المركزية التي سنتعرض إليها بالدراسة المتوسط الحسابي والوسيط والمنوال ، كما سنتعرض بالدراسة لحساب كل منهم من البيانات المفردة (الغير مبوبة) ومن البيانات المبوبة .

أولاً : المتوسط الحسابي :

أهمية المتوسط الحسابى

يعتبر المتوسط الحسابى أهم مقاييس النزعة المركزية للآتى :-

١- يعطى صورة عن المستوى العام للمجموعة لأنه يمثل الدرجة التى يتجمع حولها معظم أفراد المجموعة .

٢- المقارنة بين المجموعات .

٣- يستخدم المتوسط مع كثير من الأساليب الإحصائية الأخرى فى عمليات التحليل الإحصائى كما فى حالة t-test .

٤- تحديد موقع كل فرد فى المجموعة وذلك لأن مستوى الفرد يتحدد بانحراف درجته عن متوسط المجموعة .

وعلى الرغم من المزايا العديدة التى يتمتع بها المتوسط الحسابى من بساطة وسهولة فى الحساب ، كما أنه يأخذ جميع القيم فى الحسبان إلا أنه يعاب عليه تأثره بالقيم الشاذة أو المتطرفة Outliers or Extreme Observations .

طرق حساب المتوسط الحسابى :-

أولاً : فى حالة البيانات الخام :-

المتوسط الحسابى لمجموعة من القيم الخام أو الدرجات الخام Row Scores هو عبارة عن مجموع هذه الدرجات مقسوماً على عددها ، ويعبر عن ذلك إحصائياً بالمعادلة الآتية :-

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال : أجرى اختبار لقياس الذكاء على عينة حجمها (10) أفراد وكانت درجاتهم (75, 105, 60, 100, 90, 85, 70, 110, 120, 80) . والمطلوب حساب المتوسط الحسابى ؟

الحل : لحساب المتوسط الحسابى نستخدم المعادلة :-

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{895}{10} = 89.5$$

حالـــة خاصة : إذا كان لدينا عدد كبير نسبياً من الأفراد ، وكان العدد الكلى للدرجات صغير نسبياً فإننا فى هذه الحالة نتعامل مع بينات عددية متكررة . ولحساب المتوسط الحسابى فى هذه الحالة نستخدم المعادلة :-

$$\overline{x} = \frac{\sum_{i=1}^{n} f x_i}{\sum f}$$

أى يساوى المتوسط الحسابى مجموع حاصل ضرب الدرجات فى التكرارات المناظرة لها مقسوماً على مجموع التكرارات .

مثال : أجرى اختبار لقياس تذكر المفاهيم الرياضية الأساسية بالمرحلة الإعدادية على عينة حجمهـــا (16) تلميذاً وكانت درجاتهم (6, 2, 3, 4, 6, 4, 4, 3, 4, 3, 4, 3, 4, 5, 4, 5) . والمطلوب حساب المتوسط الحسابى لتلك الدرجات ؟

الحل : نكون الجدول الآتى :-

Scores (x_i)	f	$x_i f$
2	1	2
3	4	12
4	7	28
5	2	10
6	2	12
$\sum$	16	64

ومن ثم يكون المتوسط الحسابى مساوياً :-

$$\overline{x} = \frac{\sum_{i=1}^{n} f x_i}{\sum f} = \frac{64}{16} = 4$$

ثانياً : حساب المتوسط الحسابى من التوزيعات التكرارية

أ - الطريقة المباشرة (الطريقة العامة / طريقة مراكز الفئات)

يعطى المتوسط الحسابى من المعادلة الآتية : $\overline{x} = \frac{\sum x_i f}{\sum f}$

حيث البسط يمثل مجموع حواصل ضرب مراكز الفئات والتكرارات المناظرة لها ، ويمثل المقام مجموع التكرارات .

مـــثال : أجـــرى مقـــياس للاكتساب الاجتماعى على عينة حجمها (100) من طلاب الفرقة الثالثة بكلية التربية بقنا فكانت بياناتهم كالآتى :-

I	50-59	60-69	70-79	80-89	90-99	100-109	110-119
f	6	12	15	24	18	14	11

والمطلوب حساب المتوسط الحسابى بطريقة مراكز الفئات ؟

الحل : نكون الجدول الآتى :-

Interval	f	x	xf
50-59	6	54.5	327
60-69	12	64.5	774
70-79	15	74.5	1117.5
80-89	24	84.5	2028
90-99	18	94.5	1701
100-109	14	104.5	1463
110-119	11	114.5	1259.5
$\sum$	100		8670

ويعطى المتوسط الحسابى بالطريقة العامة من العلاقة الآتية :-

$$\overline{x} = \frac{\sum x_i f}{\sum f} = \frac{8670}{100} = 86.7$$

الوسط الحسابي المرجح (الموزون) لأوساط حسابية

اذا كان لدينا مجموعات ذات إعداد مختلفة من البيانات ، وعلم الوسط الحسابي لكل مجموعة ... كيف نحصل على الوسط الحسابي للمجموعات كلها ؟؟

$$\overline{x}_w = \frac{\sum x\,w}{\sum w}$$

بافتراض ان لدينا مجموعة N1 من القيم و وسطها الحسابي X_1 ، و مجموعة N_2 من القيم و وسطها الحسابي X_2... فإن الوسط المرجح للمجموعتين =

$$\overline{X} = \frac{N_1*X_1 + N_2*X_2}{N_1 + N_2}$$

فمثلا لو أخذنا خمسة طلاب ، وسجلنا درجات هؤلاء الطلاب في مقرر الإحصاء التطبيقي ، وعدد ساعات الاستذكار في الأسبوع .

مسلسل	1	2	3	4	5	sum
x (الدرجة)	23	40	36	28	46	173
w (عدد ساعات الاستذكار)	1	3	3	2	4	

نجد أن الوسط الحسابي غير المرجح للدرجة الحاصل عليها الطالب هي :

$$\bar{x}=\frac{\sum x}{n}=\frac{23+40+36+28+46}{5}=\frac{173}{5}=346$$

وإذا أردنا أن نحسب الوسط الحسابي للدرجات x المرجحة بعدد ساعات الاستذكار w ، يتم تطبيق المعادلة التالية :

$$(\overline{w})=\frac{\sum xw}{\sum w}=\frac{23\times1+40\times3+36\times3+28\times2+46\times4}{1+3+3+2+4}$$
$$=\frac{23+120+108+56+184}{13}=\frac{491}{13}=37.769$$

وهذا الوسط المرجح أكثر دقة من الوسط الحسابي غير المرجح .

الوسيط Median

بعد معرفتنا بأهم عيوب المتوسط الحسابى من حيث تأثر قيمته بالقيم الشاذة أو المتطرفة . لذلك يفضل فى هذه الحالة أن نستخدم الوسيط . ويعرف الوسيط لمجموعة من القيم بأنه القيمة التى تقع فى منتصف هذه المجموعة بعد ترتيبها تصاعدياً أو تنازلياً .

أولاً : الوسيط للدرجات الخام

ترتب الدرجات تصاعدياً أو تنازلياً ويكون الوسيط هو الدرجة التى يكون ترتيبها $\left(\frac{N+1}{2}\right)$ وذلك فى حالة إذا كان حجم العينة (N) فردياً . أما إذا كان حجم العينة زوجياً فإن الوسيط هو متوسط الدرجتين اللتين ترتيبهما $\left(\frac{N}{2}\right)$ ، $\left(\frac{N}{2}+1\right)$.

مثال : أوجد الوسيط لدرجات موضع الضبط لعينة حجمها (9) طلاب ، ودرجاتهم كالآتى : (26, 30, 22, 27, 24, 25, 28, 29, 20)

الحل : ترتب الدرجات تصاعدياً كالآتى :-

20, 22, 24, 25, 26, 27, 28, 29, 30

فيكون الوسيط هو الدرجة التى ترتيبها 5 =2/(1+9) أى أن :-

$$Median = 26$$

مثال : البيانات الآتية تمثل درجات ثمانى طلاب على اختبار لقياس الذكاء الوجدانى : 68, 87, 76, 58, 83, 70, 62, 54 . والمطلوب حساب الوسيط ؟

الحل : الدرجات مرتبة تصاعدياً :-

54, 58, 62, 68, 70, 76, 83, 87

لذلك يكون الوسيط مساوياً :-

$$Median = \frac{(70+68)}{2} = \frac{138}{2} = 69$$

ثانياً : الوسيط للتوزيعات التكرارية

إذا كانت البيانات معروضة فى جدول توزيع تكرارى فإن الوسيط هو القيمة التى تتوسط الدرجات بعد ترتيبها تصاعدياً أو تنازلياً وعملية ترتيب الدرجات تصاعدياً أو تنازلياً فى حالة الجداول التكرارية ما هو إلا تكوين الجداول التكرارية المتجمعة Cumulative Frequency الصاعدة أو الهابطة .

ويعطى الوسيط فى حالة الجداول التكرارية المتجمعة الصاعدة من العلاقة الآتية :-

$$Med = L + \left(\frac{.5N - f_{cum}}{f_{MED}}\right) \times C$$

حيث :

L الحد الأدنى للفئة الوسيطية وهى الفئة المفترض أن يقع بها الوسيط .

f_{MED} هـو تكـرار الفـئة الوسـيطية ، f_{cum} هـو التكرار المتجمع الصاعد للفئة قبل الوسيطية ، C طول الفئة .

مـثال : أجرى اختبار لقياس سمة الخجل على عينة حجمها (100) من طلاب الفرقة الثالثة بكلية التربية بقنا فكانت بياناتهم كالآتى :-

I	50-59	60-69	70-79	80-89	90-99	100-109	110-119
f	6	12	15	24	18	14	11

والمطلوب حساب الوسيط ؟

الحل : نكون الجدول الآتى :-

Interval	f	f_{cum}
50-59	6	6
60-69	12	18
70-79	15	33
80-89	24	57
90-99	18	75
100-109	14	89
110-119	11	100
$\sum$	100	

وحيث أن ترتيب الوسيط (50 = .5N) أى أن الفئة الوسيطية (الفئة التى يقع بها الوسيط) هى (89.5 - 79.5) بالتالى يمكن حساب الوسيط من المعادلة الآتية :-

$$Med = 79.5 + \left(\frac{50-33}{24}\right)\times 10 = 79.5 + 7.08 = 86.58$$

والتفسـير السـيكولوجى للناتج يتمثل فى أن (50) طالباً يظهرون سمة الخجل وهم الطلاب الحاصـلون علـى أعلى من قيمة الوسيط على مقياس سمة الخجل أما باقى الطلاب فإنهم عاديون وأن خجلهم غير مرضى .

ومن أهم ما يتميز به الوسيط أنه لا يتأثر بالقيم الشاذة أو المتطرفة مثل المتوسط الحسـابى ، كمـا يتمـيز بـأن مجموع الانحرافات المطلقة عن الوسيط أصغر من مجموع الانحرافات المطلقة عن المتوسط الحسابى . ويعاب على الوسيط أنه لا يأخذ جميع القيم فى الحسبان حيث يأخذ فى الاعتبار القيمة أو القيمتين اللتين تقعان فى منتصف البيانات.

المنوال Mode

يعــرف المــنوال علــى أنه أكثر القيم شيوعاً أو تكراراً من غيرها . ويلاحظ أن المنوال مثل الوسيط لا يتأثر بالقيم الشاذة أو المتطرفة

أولاً : المنوال للبيانات الخام

البيانات (70,80,59, 80,65,90, 48, 55) وحيدة المنوال (Mode Unimodal = 80) ، البيانات (12, 9, 6, 8, 5, 10, 7) ليس لها منوال ، والبيانات (6, 9, 7, 6, 5, 8, 7, 9, 12, 10) لها ثلاثة مناويل Multimodal (Mode = 6, 9, 7) .

ثانياً : المنوال للتوزيعات التكرارية

يعطى المنوال فى حالة التوزيعات التكرارية من العلاقة :-

$$Mode = L + \left(\frac{f_2}{f_1 + f_2} \right) \times C$$

حيــث : L الحــد الأدنى للفئة المنوالية وهى الفئة المفترض أن يقع بها المنوال وهى التى تقــابل أكــبر تكــرار ، f_1 هو تكرار الفئة قبل المنوالية ، f_2 هو تكرار الفئة بعد الفئة المنوالية ، C طول الفئة .

مــثال : أجرى اختبار لقياس سمة الخجل على عينة حجمها (100) من طلاب الفرقة الثالثة بكلية التربية بقنا فكانت بياناتهم كالآتى :-

I	50-59	60-69	70-79	80-89	90-99	100-109	110-119
f	6	12	15	24	18	14	11

والمطلوب حساب المنوال ؟

الحــل : أكــبر تكرار فى التوزيع هو (24) وبالتالى فإن الفئة المنوالية هى الفئة (80-89) وبذلك يمكن حساب المنوال من العلاقة الآتية :-

$$Mode = 79.5 + \left(\frac{18}{15 + 18} \right) \times 10 = 79.5 + 5.45 = 83.95$$

مقاييس التشتت Measures of Variability

إن النتائج التى نخرج بها من المتوسطات الحسابية مضللة إلى حد كبير إن لم تقترن بمعامل آخر هو التشتت والدليل على ذلك الأمر أنه لو كانت لدينا مجموعتان من الأفراد طبق عليهما أحد اختبارات القدرات وكان حجم الأفراد فى كل مجموعة أربعة وكانت درجات المجموعة الأولى (50, 0, 5, 25) ودرجات المجموعة الثانية (20, 21, 21, 18) وبحساب متوسط المجموعة الأولى والثانية نجده (20) . ويعنى ذلك أن المتوسط للمجموعتين متساو رغم أن الأفراد فى المجموعة الثانية متقاربين فى درجاتهم من بعضهم البعض ومن المتوسط ، إلا أنه فى المجموعة الأولى نجد أن درجات الأفراد متباعدة عن بعضها البعض ورغماً من ذلك فإن متوسطهم مماثل لمتوسط المجموعة الثانية . ولمعرفة الوضع الحقيقى لقيم المجموعة لابد أن نقيس مدى تباعد أو تشتت القيم بعضها عن بعض ، ولا يعنى ذلك أن المتوسط لا قيمة له بل إن مقاييس التشتت تفيد فى تفسير المتوسط بل والظاهرة موضوع الدراسة ولقياس التشتت عدة أساليب منها :-

- المدى والمدى المطلق Range & Absolute Range
- الانحراف عن المتوسط Mean Devation
- نصف المدى الربيعى Semi-Interquartil Range
- الانحراف المعيارى Standard Devation
- التباين Variance

وسنخص الانحراف المعيارى Standard Devation (SD) بالدراسة لما له من فوائد فى دراسة الفروق الفردية .

الانحراف المعيارى

يعد الانحراف المعيارى من أهم مقاييس التشتت . والذى يمكن عن طريقه قياس الفروق الفردية قياساً دقيقاً . ولذلك يستخدمه الباحثون للتعرف على الانحرافات الفردية عن المتوسط الجماعى ومدى انتشار الدرجات حول المتوسط ، بل وتعتمد فكرة « بايير الذكاء على هذا الانحراف .

أولاً : الانحراف المعيارى للدرجات الخام

يعطى من خلال إحدى الصيغ الثلاثة المتضمنة فى المعادلة الآتية :-

$$SD = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n}} = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = \sqrt{\frac{\sum x_i^2}{n} - (\bar{x})^2}$$

مثال : أوجد الانحراف المعيارى والتباين للدرجات الخام الآتية :

35, 37, 20, 44, 30, 39, 33

الحل : نكون الجدول الآتى :-

(x_i)	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$
35	+1	1
37	+3	9
20	-14	196
44	+10	100
30	-4	16
39	+5	25
33	-1	1
238	0.00	348

فيكون الانحراف المعيارى مساوياً :-

$$SD = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n}} = \sqrt{\frac{348}{7}} = 7.05$$

ويكون التباين (مربع الانحراف المعيارى) مساوياً (49.7) .

ثانياً : الانحراف المعيارى للتوزيعات التكرارية

يعطى الانحراف المعيارى فى حالة البيانات المبوبة من العلاقة :-

$$SD = \sqrt{\frac{\sum f_i (x_i - \overline{x})^2}{\sum f_i}} = C\sqrt{\frac{\sum D^2 f}{\sum f} - \left(\frac{\sum Df}{\sum f}\right)^2}$$

مثال : أجرى اختبار لقياس سمة الخجل على عينة حجمها (100) من طلاب الفرقة الثالثة بكلية التربية بقنا فكانت بياناتهم كالآتى :-

I	50-59	60-69	70-79	80-89	90-99	100-109	110-119
f	6	12	15	24	18	14	11

والمطلوب حساب الانحراف المعيارى والتباين لهذه البيانات ؟

الحل : نكون الجدول الآتى :-

Interval	f	D	Df	D^2f
50-59	6	-3	-18	54
60-69	12	-2	-24	28
70-79	15	-1	-15	15
80-89	24	0	0	0
90-99	18	+1	+18	18
100-109	14	+2	+28	56
110-119	11	+3	+33	99
$\sum$	100	0.00	+22	270

ومن الجدول يكون الانحراف المعيارى مساوياً :-

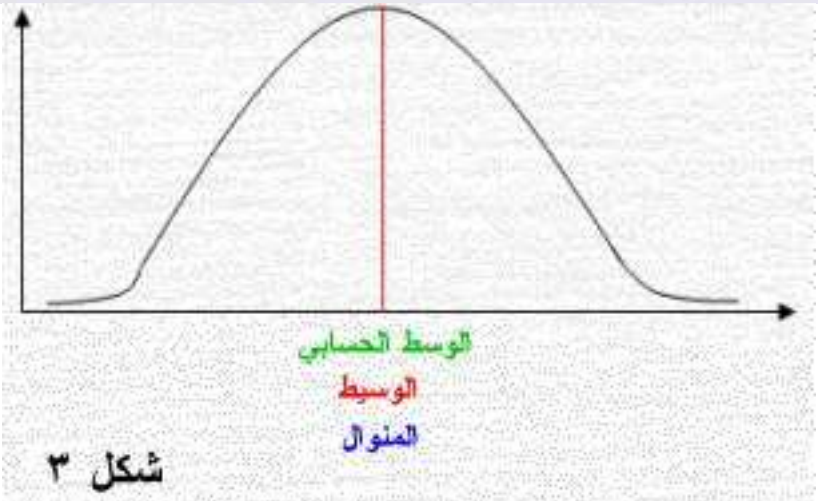
$$SD = C\sqrt{\frac{\sum D^2 f}{\sum f} - \left(\frac{\sum Df}{\sum f}\right)^2} = 10\sqrt{\frac{270}{100} - \left(\frac{22}{100}\right)^2} = 10\sqrt{2.7 - .0484} = 16.28$$

ويكون التباين (مربع الانحراف المعيارى) مساوياً (265.04) .

4-الكشف عن اعتدالية التوزيع الاعتدالي.

يعد التوزيع الطبيعي (أو ما يعرف بالتوزيع المعتدل) من التوزيعات الاحتمالية المهمة في الإحصاء وفي الدراسات التربوية والاجتماعية والإنسانية، إذ يسود اعتقاد عام مفاده أن معظم السمات والخصائص الإنسانية (درجات الذكاء، أطوال الأشخاص، التحصيل الدراسي، ... إلخ) تتوزع طبيعيًا أو تقترب من ذلك عندما يكون عدد المشاهدات كثيراً. هذا على الرغم من نسبية هذا الطرح وعدم صحته في كثير من الأحيان.

أما ما يقود إلى هذا الاعتقاد هو أن افتراض التوزيع الطبيعي يخلص الدارسين من متاعب كثيرة، إضافة إلى أن معظم الأساليب الإحصائية المستخدمة في الإجابة عن أسئلة العديد من الدراسات أو تحليل بياناتها تتطلب الاعتدالية Normality كافتراض رئيس. والمعروف كذلك أن الأساليب الإحصائية التي تستوجب توافر بعض الافتراضات حول التوزيع الاحتمالي لتوزيع البيانات تدعى الأساليب البارامترية (أو المعلمية) Parametric، أما تلك التي لا تتطلب توافر ذلك الافتراض فتعرف بالأساليب اللابارامترية (أو اللامعلمية) Non-Parametric.



شكل ٣

يمكن الكشف عن اعتدالية أي توزيع من خلال حساب معاملي الالتواء و التفلطح (التفرطح)، كما يمكن من خلال الاختبارات الإحصائية مثل اختبار χ^2 (كا^2) لحسن المطابقة ، و اختبار كولموغروف-سميرنوف **le test de** Kolmogorov-Smirnov لحسن المطابقة.

إن الكشف عن اعتدالية التوزيع خاصية هامة من خصائص الإحصاء الوصفي و يساعدنا على الاختيار بين احد الأسلوبين في عملية التحليل و الاستدلال الإحصائي ، و هما الإحصاء البارامتري و الإحصاء اللابالرامتري. فإذا كان التوزيع معتدلا فان التحاليل الإحصائية تكون وفق الإحصاء البارامتري ، أما إذا كان التوزيع غير معتدل أي حر فان التحاليل الإحصائية تكون وفق الإحصاء اللابارامتري.

و يكون توزيع البيانات معتدلا إذا كان معامل الالتواء يساوي صفر ، و يكون في نفس الوقت معامل التفلطح يساوي 3 أو قريب منه.

إن معامل الالتواء وحده لا يكفي للحكم على اعتدالية التوزيع و هذا خطا شائع في كثير من الدراسات العلمية ، فمعامل الالتواء يبين لنا فقط مدى تناظر المنحنى حول المتوسط الحسابي ام لا.

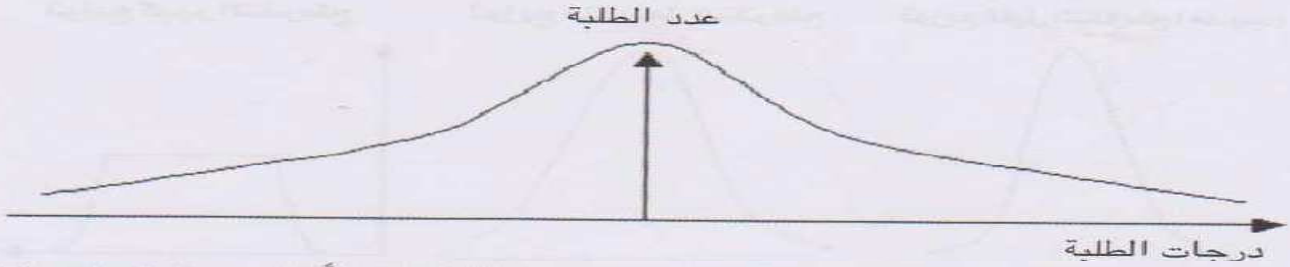
غير أن التوزيع المعتدل يتميز أيضا بخاصية التفلطح ، فقد تجد منحنيات متناظرة يكون التواءها يساوي صفر ، و لكن في نفس الوقت غير اعتدالية ، و قد تكون متناظرة فعلا من حول المتوسط الحسابي و لكنها مدببة أو مفلطحة.

يمكن أن نخلص إلى القول بان التوزيع المعتدل يتميز بخاصية التناظر حول المتوسط الحسابي و بارتفاع محدود يساوي 3 ، فهذان المعياران (الالتواء =0 و التفلطح =3) أساسيان للحكم على اعتدالية التوزيع . و كل توزيع لا يتوفر فيه المعياران معا يسمى توزيع حر .

الالتواء: يشير الالتواء إلى درجة البعد عن التماثل، ويمكن تقسيم المنحنيات إلى منحنيات ملتوية ومنحنيات متماثلة.

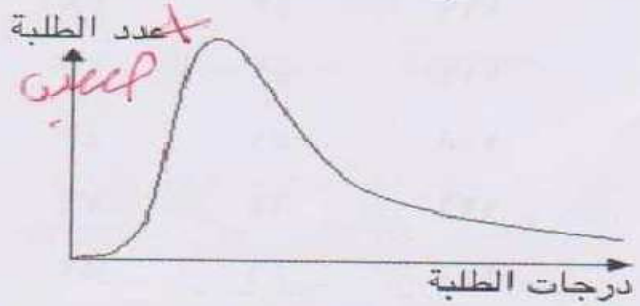
– المنحنيات المتماثلة: يعتبر التوزيع متماثلاً إذا أمكننا إقامة عمود على المحور الأفقي بحيث يقسم هذا العمود التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق. وفي الحياة العملية يوجد عدد قليل من التوزيعات المتماثلة، ولكن يوجد كثير من التوزيعات التي تكون تقريبًا متماثلة. وكمثال لتوزيع متماثل "توزيع درجات مجموعة من الطلبة على امتحان متوسط الصعوبة"، انظر الشكل التالي:

(شكل رقم ٣-٧)
توزيع متماثل

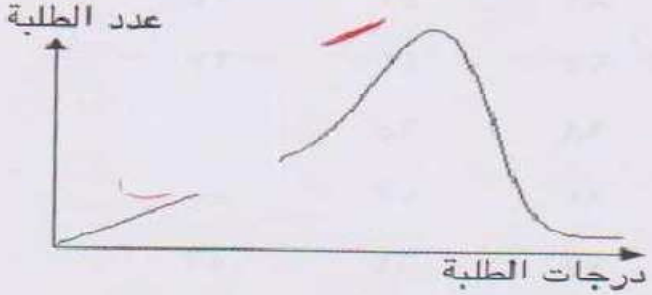


– المنحنيات الملتوية: التوزيعات التي يكون عدم تماثلها واضحًا تسمى توزيعات ملتوية. ويكون التوزيع ملتويًا إذا امتد أحد طرفيه إلى اليمين كثيرًا، أو امتد ذلك الطرف إلى اليسار كثيرًا. وكذلك يكون التوزيع ملتويًا إذا كانت القيمة العليا فيه بعيدة عن المركز أي إذا كان عاليًا من جهة ومنخفضًا من جهة ثانية. وإذا كان طرف التوزيع ممتدًا إلى اليمين (أي في الاتجاه الموجب) نقول إن التوزيع ملتو نحو اليمين أو موجب الالتواء ومثال ذلك توزيع درجات الطلبة على امتحان صعب. أما إذا كان الطرف الطويل ممتدًا نحو اليسار (أي في الاتجاه السالب) فنقول إن التوزيع ملتو نحو اليسار أو سالب الالتواء ومثال ذلك توزيع درجات الطلبة على امتحان سهل، انظر الأشكال التالية:

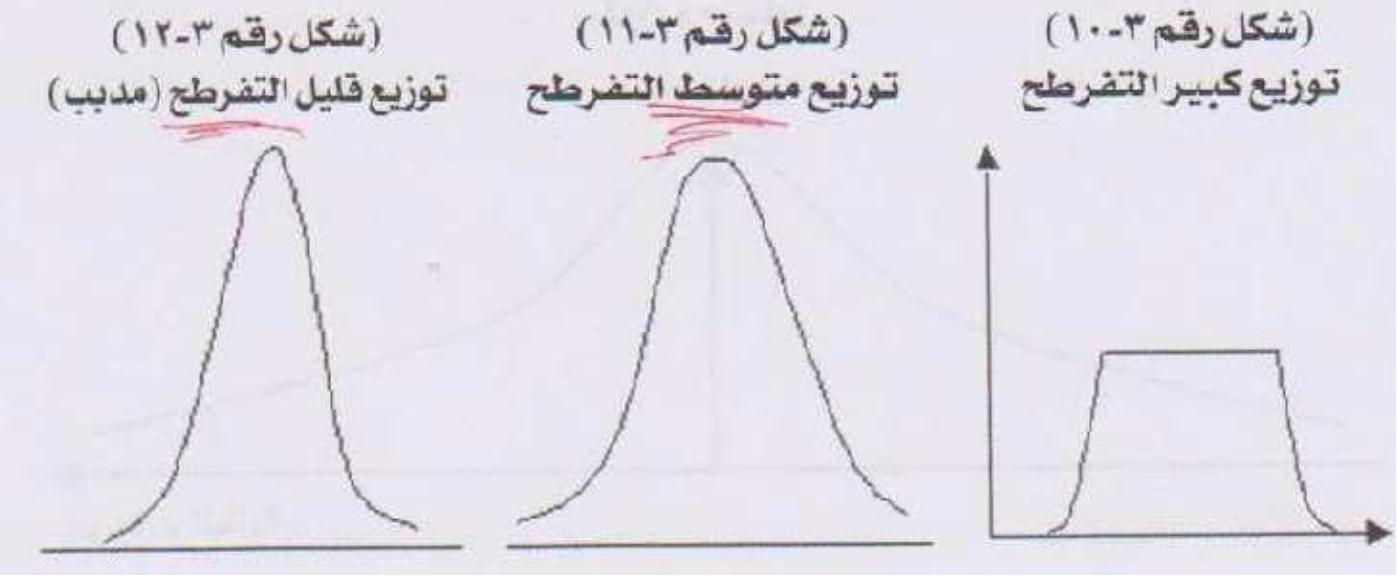
(شكل رقم ٣-٨)
توزيع ملتو نحو اليمين



(شكل رقم ٣-٩)
توزيع ملتو نحو اليسار



التفرطح: يشير التفرطح إلى درجة تدبب أو انبساط شكل التوزيع، فيمكن تقسيم المنحنيات إلى منحنيات مفرطحة ومدببة. وفيما يلى نماذج من توزيعات تختلف فى درجة التفرطح:



(أ) طريقة بيرسون لحساب الالتواء

يعطى من العلاقة الآتية :-

$$\lambda = \frac{3(Mean - Median)}{SD}$$

وفـــى ضوء هذه المعادلة نتبين أنه إذا كان المتوسط الحسابى للدرجات أعلى من الوســـيط فإن الالتواء يكون موجب والعكس صحيح أما فى حالة تساوى المتوسط والوسيط فى القيمة فإن الالتواء يصبح مساوياً الصفر ويكون توزيع الدرجات اعتدالياً Normally Distributed وهو الشرط أو الفرضية الأساسية عند معالجة البيانات بارامترياً .

حيث ان : **mean** = المتوسط الحسابي ، **Median** = الوسيط ، **SD**=الانحراف المعياري.

كما يمكن تحديد نوع التوزيع دون الرجوع إلى الرسم البياني و الاكتفاء بمقارنة مقاييس النزعة المركزية حيث يكون التوزيع اعتدالي عندما يكون : المتوسط الحسابي =الوسيط= المنوال.

(ج) طريقة العزوم لحساب معامل الالتواء

يطلق على متوسط مجموع انحرافات القيم عن متوسطها الحسابى بالعزم حول المتوسط الحسابى وتتحدد درجة هذه العزوم بالقوة المرفوعة لها هذه الانحرافات ، وتعطى العزوم الثلاثة من المعادلات الآتية :-

العزم الأول : $\mu_1 = \frac{1}{n}\sum (x - \overline{x})$

العزم الثانى : $\mu_2 = \frac{1}{n}\sum (x - \overline{x})^2$

العزم الثالث : $\mu_3 = \frac{1}{n}\sum (x - \overline{x})^3$

ويلاحظ أن العزم الأول حول المتوسط الحسابى يساوى الصفر لأن أحد المسلمات الإحصائية المعروفة لدى الباحثين أن مجموع انحرافات الدرجات عن متوسطها الحسابى يعطى صفراً ، كما أن العزم الثانى يمثل التباين ، أما العزم الثالث فيمثل التواء البيانات فى صورة مطلقة وهو يحتفظ بالإشارة سواء السالبة أو الموجبة ، ولكن لكى نحصل على مقياس نسبى وليس مطلقاً فإننا نقسم العزم الثالث على مكعب الانحراف المعيارى ، فيكون معامل الالتواء :-

$$\lambda = \frac{\mu_3}{\sigma^3}$$

ويعطى معامل التفرطح من المعادلة الآتية :-

$$\theta = \frac{\mu_4}{\sigma^4}$$

حيث يعطى العزم الرابع حول المتوسط من العلاقة :-

فى حالة الدرجات الخام :-

$$\mu_4 = \frac{1}{n}\sum (x_i - \overline{x})^4$$

حيث (x_i) هى الدرجة الخام .

وفى حالة التوزيعات التكرارية :-

$$\mu_4 = \frac{\sum f_i (x_i - \overline{x})^4}{\sum f_i}$$

حيث (x_i) هى مركز الفئة .

ويتم مقارنة معامل التفرطح بمعامل التفرطح للتوزيع الطبيعى $(\theta = 3)$. وبالتالى إذا كانت قيمة معامل التفرطح أكبر من $(\theta = 3)$ فإن ذلك يعنى أن التوزيع مدبب، أما إذا قلت قيمة معامل التفرطح عن $(\theta = 3)$ فإن ذلك يعنى أن التوزيع مفرطح .

مثال : استخدم العزوم لقياس التواء وتفرطح توزيع أعمار الطلاب بالسنوات بأحد الفصول الدراسية والتى هى على النحو التالى : 10, 14, 12, 16, 15, 11

الحل : نكون الجدول الآتى :-

x_i	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^3$	$(x_i - \overline{x})^4$
10	-3	9	-27	81
14	+1	1	+1	1
12	-1	1	-1	1
16	+3	9	+27	81
15	+2	4	+8	16
11	-2	4	-8	16
78	.000	28	.000	196

المتوسط الحسابى للدرجات :-

$$\overline{x} = \frac{78}{6} = 13$$

مثال : استخدم العزوم لقياس التواء وتفرطح توزيع أعمار الطلاب بالسنوات بأحد الفصول الدراسية والتى هى على النحو التالى : 10, 14, 12, 16, 15, 11

الحل : نكون الجدول الآتى :-

x_i	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^3$	$(x_i - \overline{x})^4$
10	-3	9	-27	81
14	+1	1	+1	1
12	-1	1	-1	1
16	+3	9	+27	81
15	+2	4	+8	16
11	-2	4	-8	16
78	.000	28	.000	196

المتوسط الحسابى للدرجات :-

$$\overline{x} = \frac{78}{6} = 13$$

التباين للدرجات :-

$$\sigma^2 = \frac{1}{n}\sum (x_i - \overline{x})^2 = \frac{28}{6} = 4.667$$

وعلى ذلك يكون العزم الأول حول المتوسط يساوى الصفر والعزم الثانى حول المتوسط يساوى تباين الدرجات والعزم الثالث حول المتوسط يعطى من :-

$$\mu_3 = \frac{1}{n}\sum (x - \overline{x})^3 = \frac{1}{6} \times .000 = .00$$

وعلى ذلك يكون التواء توزيع الدرجات :-

$$\lambda = \frac{\mu_3}{\sigma^3} = \frac{.000}{\sigma^3} = .000$$

وهذا يعنى أن توزيع أعمار الطلاب فى التجربة متماثل وغير ملتو .

والعزم الرابع حول المتوسط يساوى :-

$$\mu_4 = \frac{1}{n}\sum (x_i - \overline{x})^4 = \frac{196}{6} = 32.667$$

وبذلك يكون معامل التفرطح للبيانات :-

$$\theta = \frac{\mu_4}{\sigma^4} = \frac{32.667}{21.781} = 1.499 \simeq 1.50$$

وحيـــث أن قيمة معامل التفرطح للبيانات قلت عن $(\theta = 3)$ فذلك يعنى أن توزيع أعمار الطلاب مفرطحاً وبالتالى فإن توزيع الدرجات متماثل ولكن قمته أقل من قمة التوزيع الطبيعى .